

N 71 10540

CR 111132

RF Project 2241  
Report No. 5

FINAL REPORT

/ BY

THE OHIO STATE UNIVERSITY

RESEARCH FOUNDATION

1314 KINNEAR ROAD

COLUMBUS, OHIO 43212

# CASE FILE COPY

To NATIONAL AERONAUTICS AND SPACE ADMINISTRATION  
WASHINGTON, D.C. 20546

Grant No. NGR-36-008-040, Supplement 2

On NONPARAMETRIC RANKING AND SELECTION PROCEDURES

For the period October 1, 1968 - September 30, 1969

Submitted by Dr. M. Haseeb Rizvi  
Department of Mathematics

## NONPARAMETRIC RANKING AND SELECTION PROCEDURES

This report summarizes research performed under Grant No. 36-008-040 Supplement 2 of NASA. The objective of this grant is research and development of nonparametric methodology in the relatively new area of statistical ranking and selection procedures.

Several problems have been studied and a couple of papers written. The technical discussion of these follows.

A critical analysis of certain multiple-decision procedures based on rank sums has been carried out; these procedures have been proposed for analyzing data in a one-way layout:

$$X_{ij} = \theta_i + \varepsilon_{ij}, \quad i=1, \dots, k, \quad j=1, \dots, n,$$

where the errors  $\{\varepsilon_{ij}\}$  are independent, have the same known cumulative distribution function (cdf)  $F$  and where  $\theta = (\theta_1, \dots, \theta_k)$  is unknown. Two problems are considered:

I. Select the indices of the  $t$  largest  $\theta$ -values.

II. Select a subset containing the index of the largest  $\theta$ -value.

In problem I the experimenter sets a preassigned separation threshold  $\delta^* > 0$  and preassigned probability threshold  $P^* < 1$  and requires that the procedure he uses have the property that the probability of a correct selection is not smaller than  $P^*$  whenever the  $t$  largest  $\theta$ -values are at least  $\delta^*$  larger than the rest of the  $\theta$ -values. This problem might arise if there were  $k$  different batches of certain appliances available for use and one wanted to select the  $t$  best batches. In problem II the experimenter sets only the  $P^*$  value and requires that, with probability no smaller than  $P^*$ , the selected subset contains the index of the largest

$\theta$ -value. This problem might arise in the first stage of screening different varieties of a certain instrument; one would want to reduce the number of varieties of the instrument which are to be submitted to further tests but at the same time be reasonably sure of not eliminating any variety which has great potentiality as to its effectiveness. Bartlett and Govindarajulu (1968), Lehmann (1963), Puri and Puri (1968, 1969) and Woodworth (1965) have advanced certain procedures as solutions to problems I and II. In this paper we examine these procedures and show by means of specific examples that these procedures are in fact not solutions and should be used with great caution if they are used at all. These results are reported in a paper, written by the principal investigator (supervisor) and a co-author, entitled "Selection procedures based on ranks: counter-examples concerning least favorable configurations", which is scheduled to be published in the December, 1970 issue of the Annals of Mathematical Statistics. Five preprints of this paper have already been transmitted to NASA and the reprints will be forwarded at the time they are received from the publishers.

A selection procedure for ranking several multivariate normal populations with mean vectors  $\mu_i$  and covariance matrices  $\Sigma_i$  in terms of the Mahalanobis distances  $\theta_i = \mu_i' \Sigma_i^{-1} \mu_i$  was earlier considered by K. Alam and M. H. Rizvi in "Selection from multivariate normal populations," Annals Inst. Stat. Math., Vol. 18 (1966) pp. 307-318. This procedure, based on the natural ordering of the non-central chi-squared or non-central F random variables, has now been adapted for the optimal selection of automation systems. Tables based on rather complex numerical solutions of the involved integral equations have also been prepared. These results and the tables are given in a paper, written by the principal investigator, entitled "Optimal selection of automation systems

under multivariate normal model". The paper is being submitted for publication at this time. Preprints of this paper will soon be transmitted to NASA.

Some preliminary distribution-free results have been obtained for ranking components of the vector parameter of M-ordered densities but the decision-theoretic properties of the procedure for ranking these densities require further investigation. The problem of constructing distribution-free confidence intervals for an ordered parameter is also under study at the present time. The development of a coherent theory for combining  $\pi$  into a single formulation the requirements of a confidence interval for a certain ordered parameter and the simultaneous selection of all populations having parameters equal or larger than this ordered parameter is envisaged under a follow-on grant. It is hoped that a substantial contribution can then be made to the existing methodology of the statistical multiple-decision procedures.